

# Viral RNA pseudoknots: versatile motifs in gene expression and replication

Ian Brierley\*, Simon Pennell<sup>†</sup> and Robert J. C. Gilbert<sup>§</sup>

**Abstract** | RNA pseudoknots are structural elements found in almost all classes of RNA. First recognized in the genomes of plant viruses, they are now established as a widespread motif with diverse functions in various biological processes. This Review focuses on viral pseudoknots and their role in virus gene expression and genome replication. Although emphasis is placed on those well defined pseudoknots that are involved in unusual mechanisms of viral translational initiation and elongation, the broader roles of pseudoknots are also discussed, including comparisons with relevant cellular counterparts. The relationship between RNA pseudoknot structure and function is also addressed.

## Satellite virus

A subviral agent composed of an RNA or DNA molecule that is replicated in association with a helper virus, but encapsidated in its own encoded capsid protein.

\*Division of Virology,  
Department of Pathology,  
University of Cambridge,  
Tennis Court Road,  
Cambridge CB2 1QP, UK.

<sup>†</sup>Division of Molecular  
Structure, National Institute  
for Medical Research,  
The Ridgeway, Mill Hill,  
London NW7 1AA, UK.

<sup>§</sup>Division of Structural  
Biology, Henry Wellcome  
Building for Genomic  
Medicine, University of  
Oxford, Roosevelt Drive,  
Oxford OX3 7BN, UK.

e-mails:

ib103@mole.bio.cam.ac.uk;  
spennel@nimr.mrc.ac.uk;  
gilbert@strubi.ox.ac.uk  
doi:10.1038/nrmicro1704

The pioneering work that established pseudoknots as a genuine folding motif in RNA was carried out in the laboratories of Cornelis Pleij, Krijn Rietveld and Leendert Bosch in the early 1980s. These authors were investigating how it was possible for the 3' ends of some plant virus genomes to possess a number of the functional characteristics of transfer RNAs (tRNAs) yet lack an obvious clover-leaf secondary structure. By applying a 'pseudoknot building principle', it became clear how these viral RNAs could fold into L-shaped structures resembling tRNAs<sup>1,2</sup>. A seminal paper from the same authors<sup>3</sup> subsequently defined the general principles of pseudoknot folding and provided the first examples of pseudoknots in other RNAs — the central pseudoknot of *Escherichia coli* 16S ribosomal RNA (rRNA) and a pseudoknot present in group I introns.

Since then, many more pseudoknots have been discovered, and they are associated with a remarkably diverse range of biological activities (**Supplementary information S1, S2** (tables); reviewed in REFS 4–11). They are especially associated with key roles in the replication cycles of numerous animal and plant viruses, including, in humans, the flavivirus hepatitis C virus (HCV)<sup>12</sup>, the coronavirus responsible for severe acute respiratory syndrome (SARS-CoV)<sup>13</sup>, the oncogenic retrovirus T-cell lymphotropic virus types I and II<sup>14</sup>, and certain strains of HIV<sup>15</sup>.

The function of a viral pseudoknot is linked logically to its location in the genome (FIG. 1; **Supplementary information S1** (table)). So, in non-coding regions (NCRs) of positive-strand RNA viruses (in which the

genomic RNA serves as the mRNA template for translation and then as a template for replication), pseudoknots act in the regulation of initiation of protein synthesis and in template recognition by the viral replicase. By contrast, in coding regions they modulate the elongation and termination steps of translation. Fewer pseudoknots have been documented in the mRNAs of viruses with DNA genomes, but several DNA bacteriophage mRNAs are known to encode pseudoknots<sup>16</sup>. In the 5' NCRs, these motifs have roles in the regulation of translation initiation, whereas in the coding region they affect translation elongation. Pseudoknots have also been described in the catalytic RNAs of some RNA satellite viruses, where they have a role in genome replication<sup>17</sup>.

Here, we review selected, well characterized examples of pseudoknots in virus genomes — with an emphasis on structure–function relationships — highlighting recent advances in our understanding of pseudoknot conformation at high resolution and exploiting, where relevant, our improved knowledge of ribosome architecture.

## What is an RNA pseudoknot?

As defined originally<sup>3</sup>, a pseudoknot is a structure formed upon base-pairing of a single-stranded region of RNA in the loop of a hairpin to a stretch of complementary nucleotides elsewhere in the RNA chain (FIG. 2). Such pseudoknots, referred to as hairpin type (H-type) pseudoknots, have two base-paired stem regions (S1 and S2) and, depending on the number of loop bases that participate in the pseudoknotting interaction<sup>18</sup>, two or three single-stranded loops (L1, L2 and L3). In most (>85%;

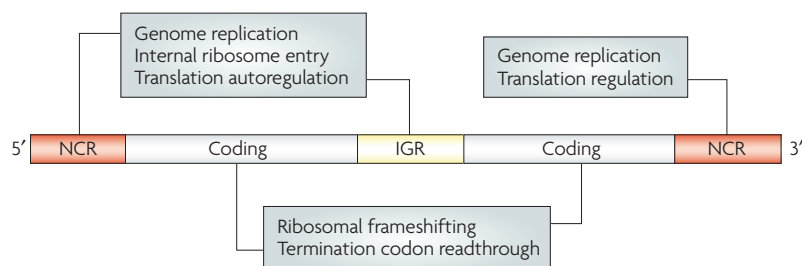
**Peptidyl (P) site**

The site on the small ribosomal subunit that holds the tRNA molecule that is linked to the growing end of the polypeptide chain.

REFS 19,20) H-type pseudoknots, L2 is absent or very short, and the base-paired stems stack coaxially to form a quasi-continuous helix. In these structures, L1 spans S2 and crosses the deep groove of the helix, whereas L3 spans S1 and crosses the shallow groove (FIG. 2).

The geometry of the pseudoknot is such that when S2 is six or seven base pairs in length, L1 can be as short as a single nucleotide<sup>3,21</sup>. However, in some pseudoknots, the loops are much longer and include their own secondary structure elements. Pseudoknots are also formed upon base-pairing of single-stranded bulge (B), interior (I) and multibranch (M) loops with complementary regions elsewhere in the RNA (which themselves can be constrained in a secondary structure, for example in intramolecular hairpin-loop-hairpin-loop (H-H or so-called kissing loop) interactions)<sup>22</sup>. However, unless all of the loop nucleotides are paired, these pseudoknots are generally considered as H-type pseudoknots, as the additional base-pairing interaction is viewed as a substructure within the loop. For this reason, the B, I and M nomenclature (as well as H-H) is not extensively used, and most structures are referred to as H-type pseudoknots and often simply as pseudoknots. An additional issue regarding nomenclature is loop numbering. In the early pseudoknot literature, most examples did not possess unpaired residues between the two stems, so the convention was to name the groove-spanning loops L1 and L2 (now L1 and L3). Here, we have opted for the L1, L2 and L3 nomenclature, which is more generally applicable.

As will be discussed in more detail below, the biological properties of pseudoknots are intimately linked to their structural features<sup>4-7,18,23</sup>. For example, the geometry of the junction between the stems and the interactions that can occur between the constituent loops and stems are often of great functional relevance<sup>21,24-27</sup>. Indeed, for many viral pseudoknots, much of the primary sequence is unimportant for function, as long as the conformation and overall stability of the structure is maintained. Where precise nucleotide-sequence requirements have been identified, this is likely to reflect a specific structural necessity, although additional roles might be possible (for example, base-specific recognition by proteins).

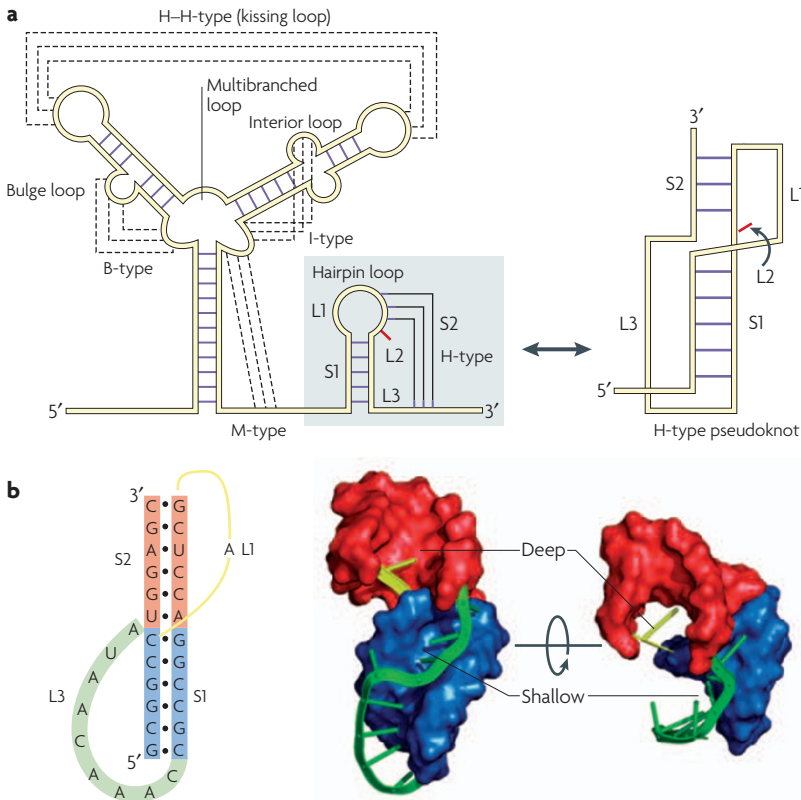


**Figure 1 | RNA pseudoknots in virus gene expression.** A schematic of a generic RNA virus genome is shown. Viral pseudoknots have been described in the 5' non-coding region (NCR), the coding region, the intergenic region (IGR) and the 3' NCR, where they function in various steps of the replication cycle. Although the majority of examples are from positive-strand RNA viruses, pseudoknots also have a role in the replication cycles of certain DNA viruses, satellite RNA viruses and viroids. For simplicity, viral pseudoknots involved in long-range interactions (including virus genome circularization) or possessing catalytic activity are not shown, but are discussed in the text.

**Pseudoknots and translation**

**Internal ribosome entry.** Most eukaryotic cellular mRNAs are translated in a cap-dependent manner, with the 40S subunit and associated initiation factors scanning along the mRNA until the start codon (AUG) is reached<sup>28</sup>. Efficient translation also requires mRNA circularization, which is brought about by the interaction of the 3'-end poly(A) tail-binding protein (PABP) with initiation factor 4E (eIF4E), bound to the 5' cap<sup>29</sup>. However, the genomes of many positive-strand RNA viruses often lack a cap, a poly(A) tail or both, and translation initiation involves non-standard mechanisms<sup>30,31</sup>.

One such example is cap-independent internal ribosome entry, in which the ribosome is recruited internally to a structured region of the mRNA (the internal ribosome entry site (IRES), usually located in the 5' NCR) and often directly to the start codon. Pseudoknots have been identified in a number of IRESs, and their function is best exemplified in the IRESs of the flavivirus HCV and the dicistrovirus cricket paralysis virus (CrPV) (FIG. 3). Unlike the mammalian picornaviral IRESs, which retain a requirement for many or most canonical initiation factors, the mammalian 40S subunit can associate with the HCV IRES in the absence of translation initiation factors, although the formation of the 48S complex (with the initiation codon locked into the mRNA-binding cleft of the small subunit) requires the participation of the eIF2-GTP-Met-tRNA<sub>i</sub> ternary complex and eIF3 (reviewed in REF. 32). The HCV IRES forms a defined secondary structure that contains two major hairpins (domains 2 and 3) and an essential pseudoknot structure<sup>12</sup> at the base of domain 3 (domain 3e/f) (FIG. 3). IRES function requires domains 2 and 3, but initial binding of the 40S subunit is mediated principally by the basal region of domain 3, including stem-loop 3d, with a modest contribution from the pseudoknot<sup>33,34</sup>. Comparisons of HCV IRES-40S (REF. 35) and IRES-80S (REF. 36) complexes have revealed that the overall appearance of the IRES is similar in the two complexes<sup>36</sup>. In the 80S complex, the pseudoknot corresponds to an L-shaped density located at the mRNA exit channel in the vicinity of ribosomal protein S5 (rpS5) (FIG. 3; **Supplementary information S3** (figure)). The assignment of the pseudoknot to the L-shaped density is based on molecular modelling and is consistent with recent crosslinking data<sup>37</sup>. The three-dimensional (3D) structure of the HCV pseudoknot is not available, but in essence it is an H-type pseudoknot in which L1 is long and highly structured, and includes the entirety of domains 3a-3e. Why a pseudoknot is present at this key location of the IRES is unclear, but it might be linked to a capacity to bind rpS5 (REF. 37). This protein is present at the mRNA exit channel and also associates with IRES domain 2 (REFS 36,38). Boehringer and colleagues propose that the HCV IRES domains function synergistically to position the AUG into the ribosomal peptidyl (P) site, coupled to movement of the pseudoknot<sup>36</sup>. In this model, a conformational change of the four-way junction (which includes domains 3a and 3c; FIG. 3) pivoted around domain 3d is transmitted to the pseudoknot,



**Figure 2 | RNA pseudoknot structure. a** | Various structural motifs have been described in RNA<sup>149</sup>. Orthodox secondary structures consist of base-paired regions (stems) connected by single-stranded loops at stem termini (hairpin loop), or in the body of a stem (bulge (B) or interior (I) loop) or at the junction of several stems (multibranch (M) loop). Pseudoknots are considered as a tertiary structure and form when bases in a loop pair with a single-stranded region elsewhere. The hairpin type (H-type) pseudoknot is by far the most common, and this tertiary interaction involves bases in the loop of a hairpin loop. The resultant structure contains two stem regions, S1 and S2, connected by single-stranded loops. In many cases, no unpaired bases are present between the two stems (L2 is zero), and the stems stack coaxially to give a quasi-continuous helix. **b** | The secondary structure of the pseudoknot of the ribosomal frameshifting signal of simian retrovirus 1 (SRV-1) is shown alongside three dimensional views of the nuclear magnetic resonance model<sup>74</sup>. The stems are shown as surface representations and the loops as ribbons (all structural images were prepared using PyMol). The polarity and handedness of the double helix leads to inequivalence of the loops, with L1 (yellow) crossing the deep groove and L3 (green) crossing the shallow groove. S1 is blue, S2 is red, L1 is yellow and L3 is green. L2 is not present in the example shown.

which moves towards the mRNA exit channel, positioning the AUG correctly into the ribosomal P site and allowing subunit joining. The movement could be potentiated by eIF3 binding, as this multisubunit initiation factor makes intimate contacts with the HCV IRES<sup>39</sup>.

Pseudoknots also have an essential role in the function of the intergenic region (IGR) IRES of CrPV<sup>40</sup> and other dicistroviruses<sup>41</sup>. This remarkable IRES, only ~200 nucleotides in length, has been described as an RNA-based translation factor<sup>42</sup> because it recruits ribosomes and activates translation without the involvement of initiation factors or initiator tRNA. The ribosomal 40S and 60S subunits bind directly to the IRES<sup>43</sup>, which then occupies the ribosomal intersubunit space of the 80S complex and interacts with key components that form the ribosomal aminoacyl (A), P and exit (E) sites.

**Aminoacyl (A) site**

The site on the small ribosomal subunit that accepts and proofreads incoming aminoacyl tRNAs prior to peptidyl transfer.

**Exit (E) site**

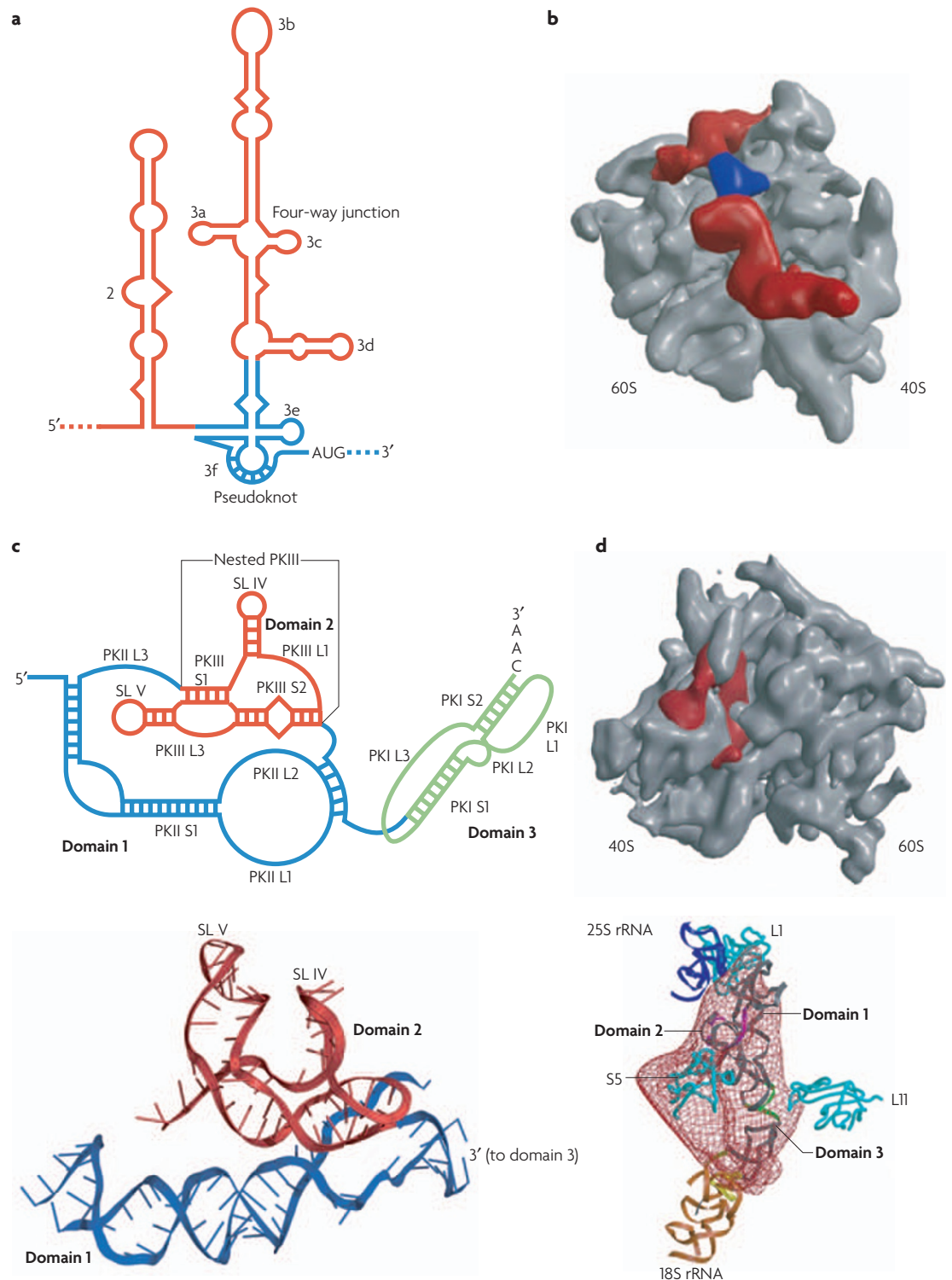
The site on the small ribosomal subunit through which tRNAs pass after they have donated their amino acid to the growing polypeptide chain.

The three defined domains of this IRES have distinct functional tasks (FIG. 3). Domain 1 contributes to interactions with the 60S subunit in the E- and P-site regions, and domain 2 interacts with the 40S subunit at the E site. Domain 3, which is located predominantly between the A and P sites and is in a similar orientation to ribosome-bound tRNAs, places its most 3' nucleotide triplet (GCU) into the decoding region of the A site<sup>11</sup>. Here, it binds the anticodon of tRNA<sup>Ala</sup>, which is brought to the ribosome as part of the ternary complex (elongation factor 1A (eEF1A)–GTP–tRNA<sup>Ala</sup>). Subsequently, tRNA<sup>Ala</sup> is pseudotranslocated (without peptide-bond formation) by eEF2 into the P site, allowing delivery of the next tRNA into the A site and authentic elongation to begin<sup>41,43,44</sup>.

Modelling and structural analysis of the CrPV IRES and IRESs from related viruses, including *Plautia stali* intestine virus (PSIV)<sup>26,27,45,46</sup>, has revealed that the structure is dominated by three H-type pseudoknots (FIG. 3), one per domain. Pseudoknot PKI, which essentially forms all of domain 3, is characterized by the possession of an AU-rich S1 and short loops. The folding of the rest of the IRES is dominated by interactions between pseudoknots PKII and PKIII (FIG. 3). The pseudoknot of domain 2, PKIII, is a nested pseudoknot in that it is entirely contained within L3 of the pseudoknot of domain 1, PKII. Cryo-electron microscopy (cryo-EM) and X-ray crystallography studies<sup>26,27</sup> (reviewed in REF. 11) have indicated a complex folding strategy that forces two small hairpins of PKIII, present as substructures in L1 (stem-loop (SL) IV) and L3 (SL V), to project from a central core and emerge on the same side of the structure to make vital interactions with rpS5 at the E site<sup>46–48</sup>. Pivotal to this folding strategy are pseudoknot loop–helix interactions. In PKIII, for example, an L1–S2 major-groove interaction positions SL IV, and a second interdomain interaction occurs between the minor groove of S2 and four L1 bases of the adjacent pseudoknot PKII, which stabilizes the core. The geometry of the overall fold is such that S2 of PKII stacks on S2 of PKIII to create a wedge-shaped section that occupies the mRNA channel and directs PKI into the decoding site.

Recruitment of the first aminoacyl tRNA to the ribosome requires the participation of eEF2 (REF. 49). The activation of eEF2 has been linked to an IRES-mediated destabilization of a conserved cellular pseudoknot that is present in helix 18 of the small subunit rRNA (the so-called 530-loop pseudoknot; **Supplementary information S2** (table)). This region of rRNA is involved in the enhancement of translational accuracy and tRNA binding<sup>50</sup>, and its destabilization (probably by the pseudoknot of domain 3) is likely to be crucial to the IRES mechanism.

The structure of the IGR IRESs illustrates how pseudoknots can be used to direct the global folding of an RNA sequence. The pseudoknot motif naturally provides the potential for coaxial stacking of constituent helices, but also the opportunity for additional stacking with helices present in constituent loops. This allows longer helical domains to be generated, a common feature in the organization of global RNA structures<sup>51</sup>. If the pseudoknot is



**Figure 3 | Pseudoknots and internal ribosome entry.** **a** | A secondary structure representation of the hepatitis C virus (HCV) internal ribosome entry site (IRES) with the pseudoknot shown in blue. **b** | A surface representation of the human 80S ribosome (grey) in complex with the HCV IRES (red) derived from the cryo-electron microscopy (cryo-EM) structure<sup>36</sup>. Density corresponding to the pseudoknot is indicated in blue. **c** | A secondary structure representation of the *Plautia stali* intestine virus (PSIV) IRES is shown above a ribbon representation of the RNA (domains 1 and 2) derived from the crystal structure<sup>27</sup>. Domain 3 remains to be solved. The secondary structure of the cricket paralysis virus (CrPV) IRES (not shown) is similar. **d** | A surface representation of the yeast 80S ribosome (grey) is shown in complex with the CrPV IRES (red) derived from the cryo-EM structure<sup>26</sup>. Below is a fit of the density to the modelled CrPV IRES showing the interactions that occur between the various domains and ribosomal components. Ribosomal proteins are cyan, 25S ribosomal RNA (rRNA) is purple and 18S rRNA is brown.

itself nested in another pseudoknot, additional helical stacking possibilities are created. Superimposed on this is the capacity of the single-stranded loops to interact with constituent stems to add stability, or with other regions of RNA to promote packing of adjacent helices. The compact and complex fold of the IGR IRESs brought about by the nested pseudoknots is a strategy similar to that used to fold certain ribozyme cores, discussed in more detail below. There is no known ribozyme activity associated with IRESs, however, which indicates that this is a general folding strategy that can be used to satisfy different mechanistic requirements.

**Autoregulation.** One of the first viral pseudoknots to be described<sup>16</sup> is encoded by T-even bacteriophages (such as T2, T4 and T6) and functions in translational autoregulation of the gene 32 protein (gp32), a single-stranded DNA-binding protein that mainly functions in replication of the viral double-stranded genomic DNA. The pseudoknot is located some 40 nucleotides upstream of the translation start site (AUG) of the gp32 mRNA and acts as a specific binding site for gp32 itself. At low protein concentrations, pseudoknot-bound gp32 does not overlap the ribosome-binding site, but as protein levels increase, cooperative binding of multiple copies occurs, nucleated at the pseudoknot-bound gp32 (REF. 52). This assembly blocks access to the Shine–Dalgarno sequence and so represses translation. Unfortunately, molecular details of the gp32–pseudoknot interaction are lacking, but the structure of the pseudoknot isolated from bacteriophage T2 has been solved by nuclear magnetic resonance (NMR)<sup>53</sup>. It is a classic H-type pseudoknot with short loops and coaxially stacked stems, albeit with S2 rotated by  $\sim 18^\circ$  with respect to S1 to relieve close phosphate–phosphate contacts at the junction while preserving the stabilizing effects of base stacking. Although the T2 pseudoknot possesses only a single nucleotide in L1 (an A), this is stereochemically feasible as the distance between the two phosphates across the deep groove of A-form RNA reaches a minimum when six or seven base pairs of S2 are bridged<sup>52,21</sup>.

Although the gene 32 system represents the only known viral example of pseudoknot involvement in the autoregulation of translation initiation, there are related examples in cellular mRNAs (**Supplementary information S2** (table)). For example, translational repression of the ribosomal S4  $\alpha$  mRNA operon<sup>54,55</sup> and autoregulation of ribosomal protein S15 synthesis<sup>56</sup> requires specific binding of the respective proteins to pseudoknots in the 5' untranslated region (UTR) of their own mRNAs.

**Frameshifting.** RNA pseudoknots in coding regions are principally associated with sites of programmed  $-1$  ribosomal frameshifting. This is a translational mechanism used by many viruses to coordinately express two proteins from a single mRNA at a defined ratio<sup>7,57</sup>. During elongation, ribosomes decode the mRNA in triplet steps and the reading frame is accurately maintained. However, in frameshifting, the ribosome is forced to shift one nucleotide backwards into an overlapping reading frame and translate an entirely new sequence

of amino acids (FIG. 4). In retroviruses, frameshifting at the overlap of the *gag* and *pol* open-reading frames (ORFs) allows expression of the viral Gag–Pol polyprotein and sets a defined cytoplasmic Gag:Gag–Pol ratio that is optimized for virion assembly and packaging of reverse transcriptase<sup>58</sup>. In other RNA viruses, frameshifting allows expression of RNA-dependent RNA polymerases (RdRps)<sup>59</sup>. Maintaining a precise efficiency of frameshifting has been shown to be crucial to the replication of HIV-1 (REF. 60) and the retrovirus-like double-stranded RNA virus of yeast, L-A<sup>61</sup>. Similarly, in other RNA viruses, changing the stoichiometry of non-frameshifted and frameshifted products is also likely to be detrimental. In SARS-CoV, for example, components of the viral replication machinery present in the viral polyproteins pp1a and pp1ab (which are expressed by frameshifting) are predicted to form a heterodimer with a stoichiometry of  $\sim 8:1$  (REFS 62,63), a ratio that is consistent with the natural level of frameshifting<sup>64,65</sup>. For these reasons, frameshifting has emerged as a potential target for antiviral therapeutics.

A typical frameshift signal has two essential elements: a heptanucleotide 'slippery' sequence, at which the ribosome-bound tRNAs slip into the  $-1$  frame, and an adjacent mRNA secondary structure, most often an H-type pseudoknot<sup>14,66</sup>, that stimulates this slippage process (FIG. 4). How the pseudoknot promotes frameshifting is not completely understood, but the mechanism is likely to be linked to the helicase activity of the ribosome, with the pseudoknot presenting an unusual topology that resists unwinding<sup>47,67–71</sup>. Takyar and colleagues have shown that the prokaryotic 70S ribosome can itself act as a helicase to unwind mRNA secondary structures before decoding<sup>70</sup>, with the active site located between the head and shoulder of the 30S subunit. Prokaryotic ribosomal proteins S3, S4 and S5 that line the mRNA entry tunnel are implicated in the helicase activity<sup>70</sup>. Recent cryo-EM images of mammalian 80S ribosomes paused at a coronavirus frameshift-promoting pseudoknot revealed that the ribosomes become stalled during the translocation phase of the elongation cycle with eEF2 bound and in the act of transferring peptidyl-tRNA from the A site to the P site<sup>71</sup>. Furthermore, the P site tRNA is structurally distorted and in a spring-like conformation. Consistent with its proposed function, the pseudoknot is present at the mRNA entry channel in close proximity to the proteins (rpS3, rpS9 and rpS2) that are likely to form important elements of the mammalian 80S helicase (FIG. 4). Namy and colleagues<sup>71</sup> suggest a mechanical model of frameshifting in which the pseudoknot resists unwinding by the helicase, compromising translocation by putting tension on the mRNA, leading to bending of the tRNA anticodon and, ultimately, repositioning the tRNA on the slippery sequence in the  $-1$  reading frame. The pseudoknot is also in close proximity to the ubiquitous ribosomal regulatory protein RACK1 (receptor for activated C kinase 1)<sup>72</sup>. It is not known whether this protein, or the recruited kinase, has a role in frameshifting, but this could in principle provide a route to regulation of frameshifting during virus infection.

#### Ribozyme

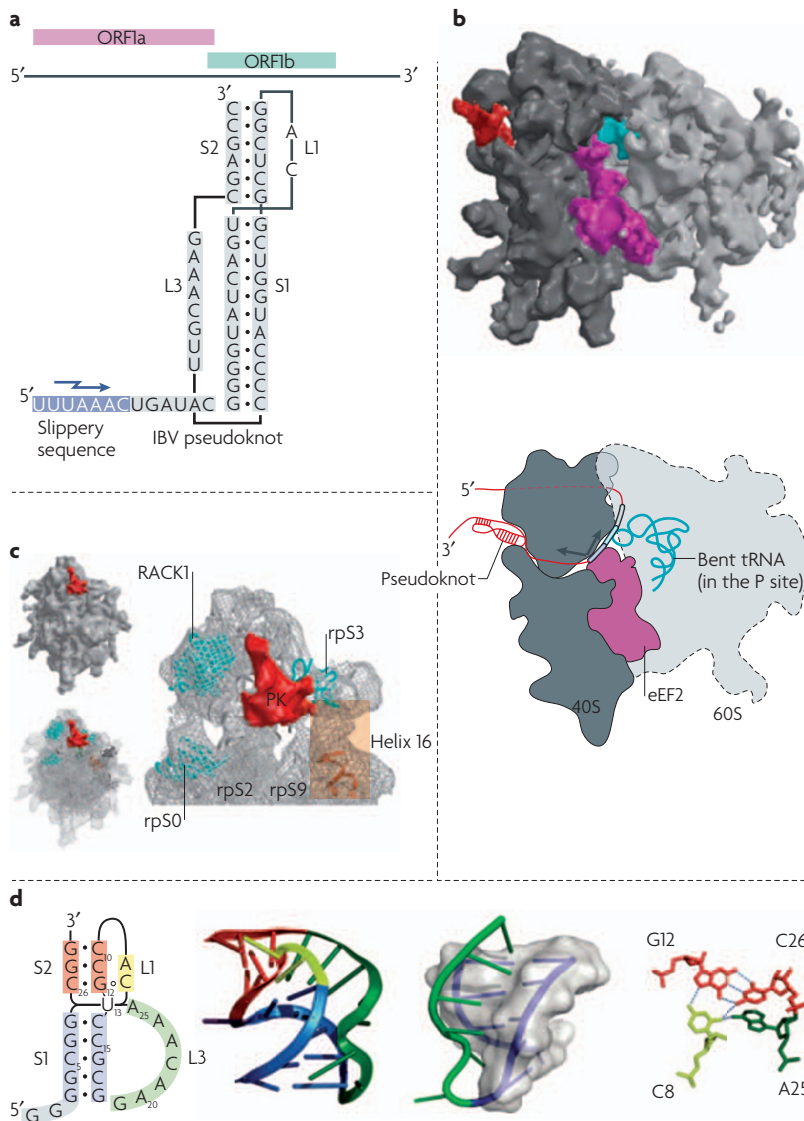
An enzyme that has an RNA as its catalytic component.

#### Shine–Dalgarno sequence

(AGGAGG). This sequence is located 5' of the AUG codon on bacterial mRNAs and functions as the signal for the initiation of protein synthesis.

#### Transfer messenger RNA

(tmRNA). An RNA that possesses both tRNA and mRNA characteristics that functions in the recognition and rescue of ribosomes stalled on aberrant mRNAs, the disposal of the causative defective mRNAs and the promotion of the degradation of ribosome-associated protein fragments.



**Figure 4 | Pseudoknots and ribosomal frameshifting. a** | The overlapping coding sequences open reading frame 1a (ORF1a) and ORF1b of the genome of the coronavirus infectious bronchitis virus (IBV) are shown above the minimal frameshift-promoting sequences of this virus. The pseudoknot promotes frameshifting at the slippery sequence, indicated by a jagged arrow. **b** | A representation of the stalled, pseudoknot-engaged rabbit 80S ribosome is shown derived from the cryo-electron microscopy structure<sup>71</sup>. The 60S subunit is light grey and the 40S subunit is dark grey. The peptidyl (P)-site transfer RNA (tRNA) stalled in the complex is coloured turquoise, the eukaryotic translocase, elongation factor 2 (eEF2), is purple and the pseudoknot structure is red. Below is a schematic of the stalled ribosome. Engagement with the pseudoknot generates a frameshifting intermediate in which the ribosome is stalled during translocation with eEF2 bound, generating tension in the mRNA that bends the P-site tRNA in a (+) sense direction. As a result, the anticodon–codon interaction breaks over the slippery sequence, allowing a spring-like relaxation of the tRNA in a (–) sense direction. **c** | A close-up view of the pseudoknot (PK) in the stalled complex, with the ribosomal components (rpS0, rpS2, rpS3 and rpS9) in close proximity highlighted. **d** | The beet western yellows virus (BWYV) pseudoknot<sup>73</sup> is illustrated to show examples of features that might confound the ribosomal helicase<sup>70</sup>. Shown from left to right are: a secondary structure model of the pseudoknot, with U<sub>13</sub> drawn to indicate its extrusion from the helix; a ribbon representation of the X-ray structure; the L3–S1 (loop 3–stem 1) triple interaction, with S1 shown as a transparent surface and the helix within as a purple ribbon; and the BWYV junction quadruple interaction, with hydrogen bonds shown as dashed lines. A triplex also forms at the junction of the two stems (not highlighted). RACK1, receptor for activated C kinase 1. Panels **b** and **c** are modified with permission from REF. 71 © MacMillan Publishers Ltd.

X-ray crystallography and NMR analysis of frameshift-promoting pseudoknots, coupled with functional studies, have revealed features that could account for an intrinsic resistance to unwinding by the ribosomal helicase (FIG. 4). Chief among these is the presence of extensive minor-groove triplex interactions between S1 and the crossing L3, first observed in the pseudoknot of the luteovirus beet western yellows virus (BWYV; FIG. 4)<sup>73</sup> and subsequently in related luteoviruses and the *gag/pro* pseudoknot of simian retrovirus 1 (REF. 74). The S1–L3 RNA triplex is likely to be the first feature encountered by the elongating ribosome, and the presence of the ‘third strand’ would conceivably confound unwinding, at least temporarily.

Another feature of functional importance is the architecture of the junction between the constituent pseudoknot helices. Several non-canonical interactions have been described in and around the junction, including base triples, base quadruples and loop–loop Hoogsteen base pairing as well as distortions such as over-rotation of the stems, helical displacement and bending<sup>7,75</sup> (FIG. 4). The first NMR study of a frameshift-promoting pseudoknot, from the *gag/pro* overlap of mouse mammary tumour virus, had indeed indicated a requirement for a specific bent conformation of the pseudoknot, brought about, in part, by the presence of an intercalated unpaired adenosine residue between the two stems<sup>76</sup>. However, closely related pseudoknots with coaxially stacked stems can clearly stimulate high levels of frameshifting<sup>74,77</sup>, and it now seems likely that the specific interactions and resultant architectures of the helical junction that are required for frameshifting are strongly context dependent<sup>75</sup>. Nevertheless, there is no doubt that the junction conformation is crucial to function and might also present a kinetic or thermodynamic barrier to unwinding<sup>78,79</sup>.

Several cellular homologues of viral frameshift signals have been identified and appear to be derived from endogenous retroviruses<sup>80–82</sup> (although some recently identified candidates may have a different origin<sup>83</sup>). The pseudoknots within these genes are fundamentally similar in structure to — and retain the functionality of — their viral progenitors, and have a role in the regulation of gene expression during embryogenesis. Another cellular pseudoknot with close structural similarity to viral frameshift pseudoknots is present in bacterial transfer messenger RNA (tmRNA)<sup>84–86</sup>. This RNA, which contains four pseudoknots (PKI–IV), functions by binding to, and ultimately rescuing, ribosomes stalled on damaged mRNAs by a process termed *trans* translation<sup>86</sup>. PKI of tmRNA shows considerable similarity to the BWYV family of pseudoknots, including short stems and loops, extensive base stacking, stabilization of the stem junction by a triplex, exclusion of a uridine at the junction of the two stems and insertion of L3 into the minor groove of S1 (REF. 84). PKI functions at a different site on the ribosome than frameshift-promoting pseudoknots, however, and its precise role is uncertain. A growing number of pseudoknots have been shown to interact with the ribosome, and it is clear that they can influence function from different sites. This is illustrated

in [Supplementary information S3](#) (figure), in which the location on the ribosome of the pseudoknots involved in IRES function, frameshifting and *trans* translation are compared.

Pseudoknots can also promote +1 ribosomal frameshifting. The best characterized examples come from the cellular gene ornithine decarboxylase antizyme<sup>87</sup>. However, pseudoknot-dependent +1 frameshifting also appears to be used in the expression of certain structural proteins of the phage of Scott A (PSA) bacteriophage that infects *Listeria monocytogenes*<sup>88</sup>, although this remains the only viral example so far.

**Stop codon readthrough.** In the gammaretroviruses, which are typified by murine leukaemia virus (MuLV), *gag* and *pol* are in the same reading frame and are separated by an UAG stop codon. Here, Gag–Pol synthesis requires periodic misreading of the stop codon as a sense (Gln) codon, a process known as termination codon suppression or readthrough. Efficient readthrough is promoted by an H-type pseudoknot spaced precisely eight nucleotides downstream of the stop codon<sup>89,90</sup>. Retroviral readthrough and frameshift signals are thus similar in terms of overall organization, with the recoding site (a stop codon or a slippery sequence) separated from the pseudoknot by a spacer of similar length. There is little variation in the size of gammaretroviral pseudoknots and considerable primary sequence conservation, especially in the spacer region and L3 of the pseudoknot. Many of these bases are functionally essential<sup>89,91,92</sup>, but have not yet been linked to an explicit structural role<sup>92,93</sup>. No high-resolution structures of these pseudoknots are available as yet, and it remains to be seen whether they interact with the ribosome in a manner similar to the pseudoknots involved in frameshifting. However, it is plausible that interactions of the readthrough pseudoknot with the helicase or other ribosomal components could modulate release-factor access or activity and promote misreading of the stop codon by the near-cognate tRNA<sup>Gln</sup>. It has already been shown that sequestration of eRF1, in this case as a result of direct binding by the MuLV reverse transcriptase, can lead to increased readthrough<sup>94</sup>.

### Pseudoknots in the 3' NCR

**Translation and replication.** Several positive-strand RNA viruses harbour pseudoknots in the 3' NCRs ([Supplementary information S1](#) (table)). The best studied examples come from plant viruses and, in particular, the pseudoknot that induces the formation of the tRNA-like structure (TLS) at the end of the genome of turnip yellow mosaic virus<sup>1,2</sup> (TYMV; FIG. 5). This short H-type pseudoknot forms part of the aminoacyl acceptor arm of the TLS, stacking against an adjacent hairpin (itself the equivalent of the tRNA T-stem–loop) to generate a quasi-continuous helix that mimics the acceptor arm of tRNA and that can be adenylated and aminoacylated with Val<sup>95</sup>. NMR studies of this region of the TLS have revealed that the single-stranded loops of the pseudoknot introduce minimal distortion in the A-form helical shape of the molecule. The major groove comfortably

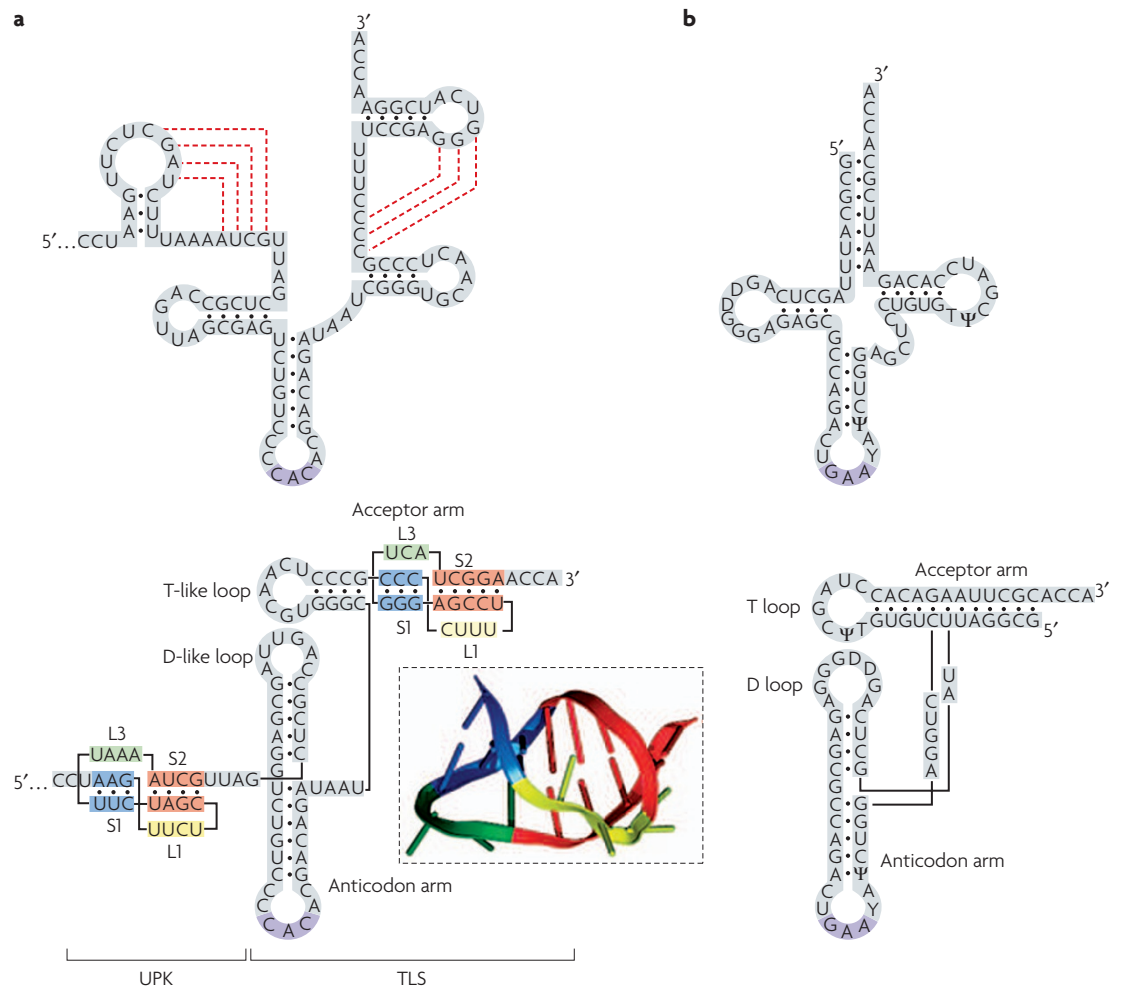
accommodates the crossing L1 residues, and L3 is closely anchored to S1 through triplex interactions centred around the adenosine residue in L3. Indeed, this study was the first to reveal the potential for interplay between pseudoknot loops and stem regions<sup>21</sup>.

The tRNA mimicry accounts for the reactivity of the 3' end of the viral genome with several enzymes that recognize tRNA. These include the CCA nucleotidyl transferase, which adds a 3' terminal A to complete the 3' CCA end of the genome upon infection; valyl tRNA synthetase, which aminoacylates the 3' end with Val; the elongation factor eEF1A, which binds to the TLS to give a viral RNA–eEF1A–GTP ternary complex; and the viral replicase p69/p206 (REFS 96,97).

An elegant relationship has been unearthed between the TLS-specific reactivities and the TYMV lifecycle. Upon entry into the cell, the 3' CCA end is completed and aminoacylated, at which point translation of the input virus genome, an obligatory step in the production of the virus replicase, is stimulated synergistically by the 5' cap and the TLS<sup>97</sup>. The stimulation of translation is maximal when the genome is aminoacylated and is linked to the formation of a viral RNA–eEF1A–GTP ternary complex<sup>97</sup>. How eEF1A enhances translation is not known, but the available evidence hints at an unexpected involvement of this elongation factor in the process of translation initiation. It has been suggested<sup>98</sup> that binding of the viral RNA ternary complex to the A-site of initiating ribosomes could stimulate initiation at the 5' end (perhaps in a manner similar to the pseudoknot of the IGR IRES domain 3 discussed above), but recent evidence argues against this specific mechanism<sup>99</sup>. Nevertheless, close proximity of the virus genome ends during translation (the circularization observed for most cellular mRNAs) could offer an explanation for the translational synergy afforded by the 5' cap and the 3' TLS–eEF1A complex.

The TLS of TYMV is also crucial in the switch between translation and replication. As has been elegantly illustrated for bacteriophage QB<sup>100</sup> and poliovirus<sup>101,102</sup>, the movement of ribosomes in the 5'→3' direction is incompatible with negative-strand RNA synthesis, in which the RNA polymerase travels 3'→5'. Thus, following an initial burst of translation, the viral mRNA must be cleared of ribosomes to allow replication from the 3' end. Negative-strand synthesis in TYMV is initiated from the second C of the 3' CCA triplet following pseudoknot-dependent binding of the replicase to the TLS<sup>103,104</sup>. It has been demonstrated recently that this reaction is inhibited by the binding of eEF1A to the valylated TLS<sup>105</sup>. So, translation is favoured until the levels of viral RdRp are sufficient to compete with eEF1A for binding to the TLS or perhaps until genomes are sequestered into vesicular sites of virus replication, which might be free of competing eEF1A and aminoacyl tRNA synthetases. The TLS might also be involved in genome packaging; in the related brome mosaic virus, viral RNAs lacking the TLS fail to assemble into virions<sup>106</sup>.

In TYMV, a second pseudoknot is present immediately upstream of the TLS, and this pseudoknot also contributes to translational enhancement, probably by



**Figure 5 | Pseudoknots and transfer RNA-like structures. a** | The 3' end of the turnip yellow mosaic virus (TYMV) genomic RNA. In the upper panel, the predicted secondary structure is shown, with pseudoknotting interactions indicated by dashed red lines. Below is a secondary structure representation of the folded molecule, showing the transfer RNA (tRNA)-like structure (TLS) and the pseudoknots in the acceptor arm and upstream of the TLS (UPK). The ribbon representation (boxed) is derived from the nuclear magnetic resonance structure of the acceptor arm pseudoknot<sup>21</sup>. **b** | For comparison, secondary structure representations of tRNA<sup>Phe</sup> are also shown. D, dihydrouridine modified bases; T, ribothymidine base.

acting as a spacing element to present the functional TLS to enzymes<sup>97</sup> (FIG. 5). The presence of such upstream pseudoknots is not uncommon and indeed, tobamo-, hordei-, furo-, pumo- and certain tymoviruses boast clusters of pseudoknots (between two and seven) in this region<sup>107</sup>, which is termed the upstream pseudoknot domain (UPD). So, in tobacco mosaic virus (TMV), for example, a total of five pseudoknots are present in the 3' UTR, three of which form the quasi-continuous double helical stalk of the UPD (5' PKIII, PKII, PKI 3') and two of which (5' PKb, PKa 3') are present in the TLS<sup>108</sup>. In the TLS, PKa forms part of the acceptor arm and PKb forms the central core, imposing the tRNA-like shape and orienting the UPD<sup>109</sup>. PKb is a B-type pseudoknot and forms a Y-shaped three-way junction that is proposed from modelling to have some structural similarity to the ribozyme of hepatitis delta virus (HDV; see below), although it displays no catalytic activity<sup>109</sup>.

Despite the complexity of folding of the TMV TLS, the UPD seems to have usurped its role, at least in terms of translational enhancement<sup>110</sup>. Although the TLS can be aminoacylated with His and can bind eEF1A<sup>22</sup>, the major player is the UPD, in which the highly conserved PKII and PKI can be crosslinked to eEF1A, independently of TLS aminoacylation<sup>111</sup>, and form an essential element of the promoter for negative-strand synthesis<sup>112</sup>. The UPD can also bind to the heat-shock chaperone HSP101, although the protein does not appear to be essential for efficient replication of TMV<sup>113</sup>. It seems that there is a general requirement for the binding of certain cellular proteins to the 3' end that does not necessarily have to be mediated solely by a TLS.

Another model system used to study the role of viral NCR pseudoknots in facilitating translation, replication and the switch between the two is tomato bushy stunt virus (TBSV)<sup>114–117</sup>. The TBSV genome is uncapped and lacks a poly(A) tail. It recruits the translation machinery initially



Box 1 | **Computational prediction of RNA pseudoknots**

RNA secondary structure prediction is not trivial and is restricted by our incomplete understanding of RNA thermodynamics and folding kinetics, and by long computer processing times<sup>137</sup>. RNAs can also adopt alternative conformations. Most programmes determine a minimum free-energy structure from the primary sequence, but cannot take into account pseudoknot topology, predominantly owing to the computational complexity<sup>138</sup>. Indeed, it has been calculated that to find a general method to search sequences for minimum free-energy pseudoknots is an intractable task<sup>139</sup>. Nevertheless, a substantial proportion (perhaps a quarter) of the pseudoknot literature deals with computational approaches to pseudoknot identification. Most commonly, a heuristic approach is taken where the search is restricted to certain pseudoknot types. However, most programmes are effective only for short sequences, as processing time can increase as the third to sixth power of sequence length, depending on the algorithm used. Another approach that has been used successfully is to perform simulations of RNA folding. These methods do not search for minimum free-energy structures and are therefore computationally more efficient. For longer sequences, a more efficient approach is to use pattern matching to perform a primary screen of a sequence database to gather those sequences with the potential to form pseudoknots of a defined type and subsequently analyse only these sequences. Multiple sequence comparisons are also of use. A combination of approaches is usually needed to have confidence in the outcome of a database search and the pseudoknots predicted within. The capacity of the programme to deal with pseudoknots containing stem regions with non-Watson–Crick pairs or bulged residues is often an issue. Despite these limitations, pseudoknot search programmes are a valuable resource. Examples of commonly used programmes and webservers are provided in TABLE 1 (see also REF. 140 for a review).

to the 3' NCR, a process requiring a highly structured 3' cap-independent translational enhancer (3' CITE), comprising a Y-shaped hairpin domain upstream of a 3'-end pseudoknot. It is thought that upon infection, the pseudoknot is folded (the closed conformation) and translation factors gather on the 3' CITE. Translation initiation complexes are then transferred to the vicinity of the 5' NCR by virtue of genome circularization, mediated, at least in part, by a kissing-loop interaction between a stem-loop in the 3' CITE and a partner close to the 5' end of the genome (SL III). Translation complexes then scan to the first AUG and begin protein synthesis. Subsequently, the pseudoknot at the 3' end of the genome is destabilized, possibly by binding of the accumulating viral RdRp, yielding an 'open' complex that is compatible with negative-strand synthesis<sup>117</sup>. Another pseudoknot (PK-TD1) in the 5' NCR is also required for efficient replication, although its mechanism of action is not fully understood.

Pseudoknots in non-coding regions have also been documented in animal RNA viruses, with the vast majority known to be essential for virus replication (Supplementary information S1 (table)). Mostly, they function in translation, genome replication or the switch between the two. However, they are not well characterized structurally (at high resolution) and few details of the molecular interactions in which they participate are available.

**Pseudoknots and ribozymes**

**Hepatitis delta virus.** HDV is a satellite RNA virus of humans that replicates in association with a helper virus, hepatitis B virus<sup>17</sup>. The circular, single-stranded RNA genome of HDV is replicated through an intermediate, the antigenome, by a double rolling-circle mechanism that requires self-cleavage by closely related genomic and antigenomic versions of a ribozyme. The HDV ribozymes have been extensively studied as a model for the mechanism of catalytic RNAs<sup>9</sup> and fold into similar structures, characterized by a nested double pseudoknot that helps form the catalytic site and brings great stability to the RNA<sup>118</sup> (Supplementary information S4 (figure)). The nested double-pseudoknot motif is also common in cellular ribozymes (Supplementary information S2 (table)), for example, the HDV-like ribozyme present in an intron of the gene that encodes human cytoplasmic polyadenylation binding protein 3 (REF. 119) and the *glmS* riboswitch ribozyme present in the 5' UTR of the gene that encodes glucosamine-6-phosphate synthase in numerous Gram-positive bacteria<sup>120,121</sup>. The intricate connectivity that results from nesting two pseudoknots within each other makes it possible for these short RNAs to adopt complex and stable 3D folds.

**A viral telomerase pseudoknot**

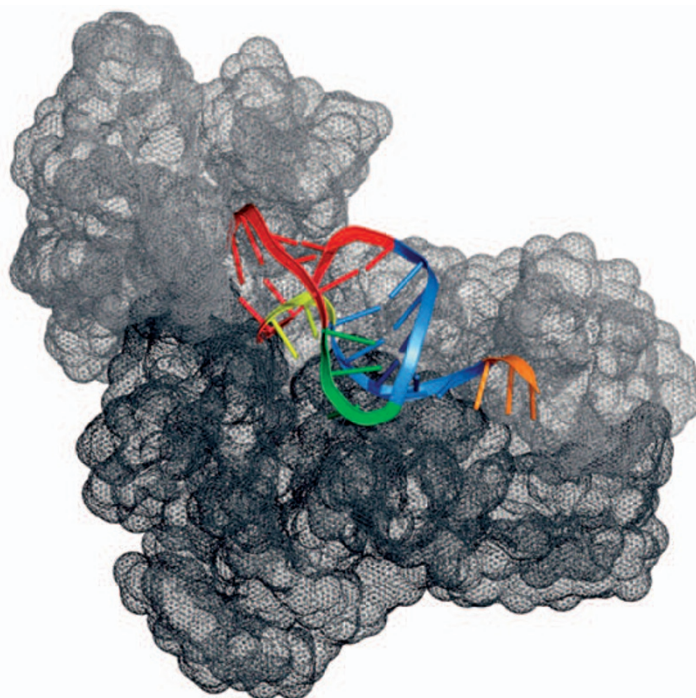
Telomerases are ribonucleoprotein complexes responsible for the maintenance of telomeres<sup>10</sup>. In addition to a specialized reverse transcriptase (TERT), an RNA component (telomerase RNA (TR)) is present and includes the RNA template for telomere addition and a highly conserved pseudoknot necessary for telomerase activity. Like frameshift-promoting pseudoknots, the TR pseudoknot has extensive triplex interactions at the junction of the two stems<sup>25</sup> (Supplementary information S5 (figure)). It has been proposed that the pseudoknot region makes contacts with the TERT<sup>122</sup> and is needed for telomere

Table 1 | **Examples of pseudoknot prediction programmes**

Task	Programme	Operating system	URL
Pattern matching	RNAmotif	Linux	<a href="http://www.scripps.edu/mb/case/casegr-sh-3.5.html">http://www.scripps.edu/mb/case/casegr-sh-3.5.html</a>
Pseudoknot prediction from short sequences	PknotsRG	Linux, OSX, Windows, Web server	<a href="http://bibiserv.techfak.uni-bielefeld.de/download/tools/pknotsrg.html">http://bibiserv.techfak.uni-bielefeld.de/download/tools/pknotsrg.html</a>
Pseudoknot prediction from short sequences	HotKnots	Linux	<a href="http://www.cs.ubc.ca/labs/beta/Software/HotKnots">http://www.cs.ubc.ca/labs/beta/Software/HotKnots</a>
Pseudoknot prediction from longer sequences	HPKNOTTER	Web server	<a href="http://bioalgorithm.life.nctu.edu.tw/HPKNOTTER">http://bioalgorithm.life.nctu.edu.tw/HPKNOTTER</a>
Pseudoknot visualization	PseudoViewer	Web server, Windows	<a href="http://pseudoviewer.inha.ac.kr">http://pseudoviewer.inha.ac.kr</a>
Pseudoknot database	PseudoBase	Web server	<a href="http://biology.leidenuniv.nl/~batenburg/PKBGetS1.html">http://biology.leidenuniv.nl/~batenburg/PKBGetS1.html</a>

## Box 2 | RNA pseudoknot aptamers

Aptamers are small nucleic acid ligands that are generated *in vitro* against various biological molecules. The majority are produced using the SELEX (selective evolution of ligands by exponential enrichment) method<sup>141</sup>. A pool of randomized RNAs is selectively enriched over repetitive rounds of target binding and subsequent sequence amplification until the surviving pool consists of sequences that bind to the target with high affinity and specificity. Initial experiments with SELEX focused on viral polymerases, confirming the sequence selectivity of T7 DNA polymerase<sup>141</sup>, and these were followed by the isolation of pseudoknotted aptamers able to bind to HIV-1 reverse transcriptase<sup>142</sup>. These aptamers have been shown to bind with picomolar affinity<sup>143</sup> and can block HIV replication in tissue culture<sup>144</sup>, inhibiting the polymerase at all stages of genome replication<sup>145</sup>. The crystal structure of aptamer-bound reverse transcriptase shows how the pseudoknot ligand (coloured strand) is bound along the cleft



between the polymerase active site and the RNase H domain<sup>146</sup> (see image; stem 1 (S1) is blue, S2 is red, loop 1 (L1) is yellow, L3 is green and additional non-pseudoknot bases are orange; dark grey is p51 and light grey is p66). The pseudoknot holds the polymerase in a 'closed' conformation, blocking its action by competitive inhibition of the primer–template binding site. Because of their specificity, there is considerable interest in the use of aptamers as therapeutics. However, aptamers cross cell membranes inefficiently owing to their hydrophilicity and are mostly restricted to extracellular applications such as the blocking of viral glycoproteins to prevent cell attachment and viral entry<sup>147</sup>. To be effective, such aptamers need to be chemically modified so as to resist host ribonucleases. SELEX experiments regularly isolate pseudoknotted ligands<sup>148</sup>, presumably reflecting the range of binding surfaces that pseudoknots can offer.

repeat-addition processivity<sup>123</sup>. Indeed, a switch between the pseudoknot and a partially unfolded form might be important in the translocation of telomerase during telomere addition<sup>124–126</sup>. High levels of telomerase activity are detected in a large range of cancers and are closely associated with the immortalization process<sup>127,128</sup>. Recently, a chicken TR homologue, vTR, has been described in the oncogenic herpesvirus Marek's disease virus<sup>129</sup>. Deletion of the two copies of vTR from the virus genome led to reduced incidence and severity of lymphoma in infected chickens, indicating that vTR has the attributes of an oncogene<sup>130</sup>. An understanding of how vTR supports lymphomagenesis in genetic and molecular terms has the potential to yield new insights into telomerase function in cancer<sup>131</sup>.

### The versatile pseudoknot

The development of sophisticated algorithms (BOX 1; TABLE 1) to scan genomic sequences for RNA structural motifs has highlighted the prevalence of the pseudoknot fold in the RNA universe, indicating evolutionary selection. A plausible explanation for this abundance is that pseudoknots offer functional versatility. Some pseudoknots (for example the BWYV frameshift pseudoknot) are extensively stabilized by both Watson–Crick and

non-Watson–Crick interactions<sup>73</sup> and are more stable than an equivalent hairpin (containing a base-paired stem that is equivalent to S1 plus S2). Thus, a very stable motif can be included in an RNA when space — either in terms of genomic coding capacity or molecular dimensions — is at a premium. At the other end of the spectrum, many pseudoknots are considerably less stable than their equivalent hairpin and could conceivably act as regulatory switches, oscillating between stem–loop and pseudoknot conformations in response to environmental signals<sup>132</sup>. Pseudoknots also offer binding sites for proteins or single-stranded loops of RNA. The often extensive intra- and intermolecular contacts that pseudoknots engage in provide many targets for such interactions. Indeed, the *in vitro* selection of RNA aptamers that bind various biomolecules often generates pseudoknotted RNAs (BOX 2). Pseudoknotting can also be the most efficient way of folding RNAs in an active conformation (for example, ribozymes).

Long-range interactions are also facilitated by this motif to organize global folding (for example, in the ribosome itself; **Supplementary information S2** (table)) and to link separate domains of RNA together. In bacteriophage Q $\beta$ , although the viral replicase is anchored at a site some 1.2 kb from the 3' end of the genome where replication

### Rolling-circle replication

A mode of replication that uses a circular molecule as a template to produce concatemers of linear molecules.

### Riboswitch

A conformational switch in an RNA molecule that is induced by a small metabolite, and that leads to a switch in gene-regulatory function.

### Aptamer

An RNA domain, either engineered or natural, that forms a precise 3D structure and selectively binds a target molecule.

initiates, an adjacent pseudoknot forms a long-range interaction that brings the 3' end to the active site of the polymerase<sup>133</sup>. Such connections also offer the potential for global regulation of gene expression. In barley yellow dwarf virus, L3 of the frameshift-promoting pseudoknot is over 4 kb in length (of a 5.6-kb genome) and links the frameshift region with the 3' end of the genome<sup>134</sup>. It has been suggested that disruption of this long-range contact is a means to regulate ribosome and replicase traffic on the viral genome<sup>135</sup>.

Especially in RNA viruses, there is often enormous selective pressure, and genomes that are optimized for viral fitness rapidly dominate. Although it is plain that some or all of the features of pseudoknots discussed above could be reproduced in other ways, it might be that this can only be achieved at considerable cost to fitness.

**Future perspectives**

Although 25 years have passed since the first pseudoknot was described<sup>1</sup>, much remains to be discovered. Foremost among the tasks ahead is to determine the true frequency of these motifs in natural RNAs. The majority of pseudoknots described in this article were identified as part of the normal process of scientific research, but computational approaches to their identification will be of increasing relevance. To maximize the proportion of genuine candidate pseudoknots identified in such screens, a better understanding of the thermodynamic parameters that govern pseudoknot formation will be invaluable. Additionally, we will need more information

about the structure and function of pseudoknots. Until recently, atomic resolution structural information has been restricted to short, often very stable, pseudoknots. Elucidating the cryo-EM and crystal structures of the IGR IRESs of CrPV<sup>26</sup> and PSIV<sup>27</sup> were huge steps forward, but similar monumental efforts will be required to solve those structures in which the pseudoknot orchestrates the folding of long and complex domains, as is found in many viral NCRs. Several viral pseudoknots remain poorly characterized, both in terms of structure and biological activity. A number of these have been shown to interact with proteins, both viral (including RdRp) and cellular, yet molecular details are lacking. Furthermore, we do not know whether pseudoknots are ever substrates for viral or cellular helicases present in infected cells; such activities could offer an extra level of control of virus gene expression and replication. The use of pseudoknots in antiviral strategies should also be considered more broadly. Antisense oligomers targeting the SARS frameshift-promoting pseudoknot have marked antiviral activity<sup>136</sup>, and pseudoknotted aptamers have been shown to block HIV replication (BOX 2).

The study of viral pseudoknots will continue to excite and challenge researchers. It is unquestionable that many more pseudoknots remain to be discovered, and it is highly likely that some of these will possess novel activities or have unprecedented functions. The field has come a long way since the early experiments of Pleij and his collaborators, yet there is still much to be discovered about these fascinating motifs.

1. Rietveld, K., Van Poelgeest, R., Pleij, C. W., Van Boom, J. H. & Bosch, L. The tRNA-like structure at the 3' terminus of turnip yellow mosaic virus RNA. Differences and similarities with canonical tRNA. *Nucleic Acids Res.* **10**, 1929–1946 (1982). **The first experimental identification of an RNA pseudoknot.**
2. Rietveld, K., Pleij, C. W. & Bosch, L. Three-dimensional models of the tRNA-like 3' termini of some plant viral RNAs. *EMBO J.* **2**, 1079–1085 (1983).
3. Pleij, C. W., Rietveld, K. & Bosch, L. A new principle of RNA folding based on pseudoknotting. *Nucleic Acids Res.* **13**, 1717–1731 (1985). **A fundamental paper describing the pseudoknot building principle, setting the rules for pseudoknot formation and providing the first examples.**
4. Pleij, C. W. Pseudoknots: a new motif in the RNA game. *Trends Biochem. Sci.* **15**, 143–147 (1990).
5. ten Dam, E., Pleij, K. & Draper, D. Structural and functional aspects of RNA pseudoknots. *Biochemistry* **31**, 11665–11676 (1992).
6. Hilbers, C. W., Michiels, P. J. & Heus, H. A. New developments in structure determination of pseudoknots. *Biopolymers* **48**, 137–153 (1998).
7. Giedroc, D. P., Theimer, C. A. & Nixon, P. L. Structure, stability and function of RNA pseudoknots involved in stimulating ribosomal frameshifting. *J. Mol. Biol.* **298**, 167–185 (2000).
8. Fechter, P., Rudinger-Thirion, J., Florentz, C. & Giege, R. Novel features in the tRNA-like world of plant viral RNAs. *Cell. Mol. Life Sci.* **58**, 1547–1561 (2001).
9. Been, M. D. Molecular biology. Versatility of self-cleaving ribozymes. *Science* **313**, 1745–1747 (2006).
10. Theimer, C. A. & Feigon, J. Structure and function of telomerase RNA. *Curr. Opin. Struct. Biol.* **16**, 307–318 (2006).
11. Hellen, C. U. Bypassing translation initiation. *Structure* **15**, 4–6 (2007).
12. Wang, C., Le, S. Y., Ali, N. & Siddiqui, A. An RNA pseudoknot is an essential structural element of the internal ribosome entry site located within the hepatitis C virus 5' noncoding region. *RNA* **1**, 526–537 (1995).
13. Thiel, V. *et al.* Mechanisms and enzymes involved in SARS coronavirus genome expression. *J. Gen. Virol.* **84**, 2305–2315 (2003).
14. ten Dam, E. B., Pleij, C. W. & Bosch, L. RNA pseudoknots: translational frameshifting and readthrough on viral RNAs. *Virus Genes* **4**, 121–136 (1990). **The first detailed description of a large number of pseudoknots that act during the elongation phase of viral protein synthesis.**
15. Baril, M., Dulude, D., Steinberg, S. V. & Brakier-Gingras, L. The frameshift stimulatory signal of human immunodeficiency virus type 1 group O is a pseudoknot. *J. Mol. Biol.* **331**, 571–583 (2003).
16. McPheeters, D. S., Stormo, G. D. & Gold, L. Autogenous regulatory site on the bacteriophage T4 gene 32 messenger RNA. *J. Mol. Biol.* **201**, 517–535 (1988). **The first description of an autoregulatory viral pseudoknot that is bound specifically by a viral protein.**
17. Taylor, J. M. Structure and replication of hepatitis delta virus RNA. *Curr. Top. Microbiol. Immunol.* **307**, 1–23 (2006).
18. Westhof, E. & Jaeger, L. RNA pseudoknots. *Curr. Opin. Struct. Biol.* **2**, 327–335 (1992).
19. van Batenburg, F. H., Gultyaev, A. P. & Pleij, C. W. PseudoBase: structural information on RNA pseudoknots. *Nucleic Acids Res.* **29**, 194–195 (2001). **A valuable resource of RNA pseudoknot information.**
20. Kim, N., Shiffeldrim, N., Gan, H. H. & Schlick, T. Candidates for novel RNA topologies. *J. Mol. Biol.* **341**, 1129–1144 (2004).
21. Kolk, M. H. *et al.* NMR structure of a classical pseudoknot: interplay of single- and double-stranded RNA. *Science* **280**, 434–438 (1998). **This work revealed the potential for pseudoknot loop–helix interactions.**
22. Mans, R. M., Pleij, C. W. & Bosch, L. tRNA-like structures. Structure, function and evolutionary significance. *Eur. J. Biochem.* **201**, 303–324 (1991).
23. Deiman, B. A. L. M. & Pleij, C. W. A. Pseudoknots: a vital feature in viral RNA. *Sem. Virol.* **8**, 166–175 (1997).
24. Kim, Y. G., Su, L., Maas, S., O'Neill, A. & Rich, A. Specific mutations in a viral RNA pseudoknot drastically change ribosomal frameshifting efficiency. *Proc. Natl Acad. Sci. USA* **96**, 14234–14239 (1999).
25. Theimer, C. A., Blois, C. A. & Feigon, J. Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Mol. Cell* **17**, 671–682 (2005).
26. Schuler, M. *et al.* Structure of the ribosome-bound cricket paralysis virus IRES RNA. *Nature Struct. Mol. Biol.* **13**, 1092–1096 (2006).
27. Pflugsten, J. S., Costantino, D. A. & Kieft, J. S. Structural basis for ribosome recruitment and manipulation by a viral IRES RNA. *Science* **314**, 1450–1454 (2006). **References 26 and 27, published simultaneously, provided a wealth of information on pseudoknot folding in large RNAs and how structural features that interact with ribosomes can be generated by such folding strategies.**
28. Marintchev, A. & Wagner, G. Translation initiation: structures, mechanisms and evolution. *Q. Rev. Biophys.* **37**, 197–284 (2004).
29. Wells, S. E., Hillner, P. E., Vale, R. D. & Sachs, A. B. Circularization of mRNA by eukaryotic translation initiation factors. *Mol. Cell* **2**, 135–140 (1998).
30. Jang, S. K. Internal initiation: IRES elements of picornaviruses and hepatitis C virus. *Virus Res.* **119**, 2–15 (2006).
31. Dreher, T. W. & Miller, W. A. Translational control in positive strand RNA plant viruses. *Virology* **344**, 185–197 (2006).
32. Fraser, C. S. & Doudna, J. A. Structural and mechanistic insights into hepatitis C viral translation initiation. *Nature Rev. Microbiol.* **5**, 29–38 (2007).
33. Kieft, J. S., Zhou, K., Jubin, R. & Doudna, J. A. Mechanism of ribosome recruitment by hepatitis C IRES RNA. *RNA* **7**, 194–206 (2001).

34. Otto, G. A. & Puglisi, J. D. The pathway of HCV IRES-mediated translation initiation. *Cell* **119**, 369–380 (2004).
35. Spahn, C. M. *et al.* Hepatitis C virus IRES RNA-induced changes in the conformation of the 40S ribosomal subunit. *Science* **291**, 1959–1962 (2001).
36. Boehringer, D., Thermann, R., Ostareck-Lederer, A., Lewis, J. D. & Stark, H. Structure of the hepatitis C virus IRES bound to the human 80S ribosome: remodeling of the HCV IRES. *Structure* **13**, 1695–1706 (2005).
37. Laletina, E. *et al.* Proteins surrounding hairpin IIIe of the hepatitis C virus internal ribosome entry site on the human 40S ribosomal subunit. *Nucleic Acids Res.* **34**, 2027–2036 (2006).
38. Fukushi, S. *et al.* Ribosomal protein S5 interacts with the internal ribosome entry site of hepatitis C virus. *J. Biol. Chem.* **276**, 20824–20826 (2001).
39. Siridechadilok, B., Fraser, C. S., Hall, R. J., Doudna, J. A. & Nogales, E. Structural roles for human translation factor eIF3 in initiation of protein synthesis. *Science* **310**, 1513–1515 (2005).
40. Wilson, J. E., Pestova, T. V., Hellen, C. U. & Sarnow, P. Initiation of protein synthesis from the A site of the ribosome. *Cell* **102**, 511–520 (2000).
41. Pflingsten, J. S., Costantino, D. A. & Kieft, J. S. Conservation and diversity among the three-dimensional folds of the dicistroviridae intergenic region IRESes. *J. Mol. Biol.* **370**, 856–869 (2007).
42. Spahn, C. M. *et al.* Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. *Cell* **118**, 465–475 (2004).
43. Pestova, T. V. & Hellen, C. U. Translation elongation after assembly of ribosomes on the cricket paralysis virus internal ribosome entry site without initiation factors or initiator tRNA. *Genes Dev.* **17**, 181–186 (2003).
44. Jan, E., Kinzy, T. G. & Sarnow, P. Divergent tRNA-like element supports initiation, elongation, and termination of protein biosynthesis. *Proc. Natl Acad. Sci. USA* **100**, 15410–15415 (2003).
45. Kanamori, Y. & Nakashima, N. A tertiary structure model of the internal ribosome entry site (IRES) for methionine-independent initiation of translation. *RNA* **7**, 266–274 (2001).
- A mutational and phylogenetic analysis of an IGR IRES that led to the realization that the IRES was composed of three pseudoknots.**
46. Jan, E. & Sarnow, P. Factorless ribosome assembly on the internal ribosome entry site of cricket paralysis virus. *J. Mol. Biol.* **324**, 898–902 (2002).
47. Yusupov, M. M. *et al.* Crystal structure of the ribosome at 5.5 Å resolution. *Science* **292**, 883–896 (2001).
48. Nishiyama, T. *et al.* Structural elements in the internal ribosome entry site of *Plautia stali* intestine virus responsible for binding with ribosomes. *Nucleic Acids Res.* **31**, 2434–2442 (2003).
49. Yamamoto, H., Nakashima, N., Ikeda, Y. & Uchiumi, T. Binding mode of the first aminoacyl-tRNA in translation initiation mediated by *Plautia stali* intestine virus internal ribosome entry site. *J. Biol. Chem.* **282**, 7770–7776 (2007).
50. Powers, T. & Noller, H. F. A functional pseudoknot in 16S ribosomal RNA. *EMBO J.* **10**, 2203–2214 (1991).
51. Batey, R. T., Rambo, R. P. & Doudna, J. A. Tertiary motifs in RNA structure and folding. *Angew Chem. Int. Ed. Engl.* **38**, 2326–2343 (1999).
52. Shamoo, Y., Tam, A., Konigsberg, W. H. & Williams, K. R. Translational repression by the bacteriophage T4 gene 32 protein involves specific recognition of an RNA pseudoknot structure. *J. Mol. Biol.* **232**, 89–104 (1993).
53. Holland, J. A., Hansen, M. R., Du, Z. & Hoffman, D. W. An examination of coaxial stacking of helical stems in a pseudoknot motif: the gene 32 messenger RNA pseudoknot of bacteriophage T2. *RNA* **5**, 257–271 (1999).
54. Tang, C. K. & Draper, D. E. Unusual mRNA pseudoknot structure is recognized by a protein translational repressor. *Cell* **57**, 531–536 (1989).
55. Schlax, P. J., Xavier, K. A., Gluck, T. C. & Draper, D. E. Translational repression of the *Escherichia coli* *a* operon mRNA: importance of an mRNA conformational switch and a ternary entrapment complex. *J. Biol. Chem.* **276**, 38494–38501 (2001).
56. Mathy, N. *et al.* Specific recognition of rpsO mRNA and 16S rRNA by *Escherichia coli* ribosomal protein S15 relies on both mimicry and site differentiation. *Mol. Microbiol.* **52**, 661–675 (2004).
57. Brierley, I. & Pennell, S. Structure and function of the stimulatory RNAs involved in programmed eukaryotic-1 ribosomal frameshifting. *Cold Spring Harb. Symp. Quant. Biol.* **66**, 233–248 (2001).
58. Brierley, I. & Dos Ramos, F. J. Programmed ribosomal frameshifting in HIV-1 and the SARS-CoV. *Virus Res.* **119**, 29–42 (2006).
59. Brierley, I. Ribosomal frameshifting viral RNAs. *J. Gen. Virol.* **76**, 1885–1892 (1995).
60. Shehu-Xhilaga, M., Crowe, S. M. & Mak, J. Maintenance of the Gag/Gag–Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity. *J. Virol.* **75**, 1834–1841 (2001).
61. Dinman, J. D. & Wickner, R. B. Ribosomal frameshifting efficiency and Gag/Gag–Pol ratio are critical for yeast M1 double-stranded RNA virus propagation. *J. Virol.* **66**, 3669–3676 (1992).
62. Zhai, Y. *et al.* Insights into SARS-CoV transcription and replication from the structure of the nsp7–nsp8 hexadecamer. *Nature Struct. Mol. Biol.* **12**, 980–986 (2005).
63. Imbert, I. *et al.* A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. *EMBO J.* **25**, 4933–4942 (2006).
64. Baranov, P. V. *et al.* Programmed ribosomal frameshifting in decoding the SARS-CoV genome. *Virology* **332**, 498–510 (2005).
65. Plant, E. P. *et al.* A three-stemmed mRNA pseudoknot in the SARS coronavirus frameshift signal. *PLoS Biol.* **3**, e172 (2005).
66. Brierley, I., Digard, P. & Inglis, S. C. Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell* **57**, 537–547 (1989).
- The first experimental verification of a role for pseudoknots in translational frameshifting.**
67. Yusupova, G. Z., Yusupov, M. M., Cate, J. H. & Noller, H. F. The path of messenger RNA through the ribosome. *Cell* **106**, 233–241 (2001).
68. Plant, E. P. *et al.* The 9-A solution: how mRNA pseudoknots promote efficient programmed-1 ribosomal frameshifting. *RNA* **9**, 168–174 (2003).
69. Plant, E. P. & Dinman, J. D. Torsional restraint: a new twist on frameshifting pseudoknots. *Nucleic Acids Res.* **33**, 1825–1833 (2005).
70. Takyar, S., Hickerson, R. P. & Noller, H. F. mRNA helicase activity of the ribosome. *Cell* **120**, 49–58 (2005).
71. Namy, O., Moran, S. J., Stuart, D. I., Gilbert, R. J. & Brierley, I. A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting. *Nature* **441**, 244–247 (2006).
- This cryo-EM study was the first experimental visualization of how a pseudoknot can compromise translational elongation.**
72. Nilsson, J., Sengupta, J., Frank, J. & Nissen, P. Regulation of eukaryotic translation by the RACK1 protein: a platform for signalling molecules on the ribosome. *EMBO Rep.* **5**, 1137–1141 (2004).
73. Su, L., Chen, L., Egli, M., Berger, J. M. & Rich, A. Minor groove RNA triplex in the crystal structure of a ribosomal frameshifting viral pseudoknot. *Nature Struct. Biol.* **6**, 285–292 (1999).
- The first crystal structure of a frameshift-promoting pseudoknot, yielding a wealth of information about pseudoknot folding.**
74. Michiels, P. J. *et al.* Solution structure of the pseudoknot of SRV-1 RNA, involved in ribosomal frameshifting. *J. Mol. Biol.* **310**, 1109–1123 (2001).
75. Cornish, P. V., Stammer, S. N. & Giedroc, D. P. The global structures of a wild-type and poorly functional plant luteoviral mRNA pseudoknot are essentially identical. *RNA* **12**, 1959–1969 (2006).
76. Shen, L. X. & Tinoco, I. Jr. The structure of an RNA pseudoknot that causes efficient frameshifting in mouse mammary tumor virus. *J. Mol. Biol.* **247**, 963–978 (1995).
77. Wang, Y. *et al.* Comparative studies of frameshifting and nonframeshifting RNA pseudoknots: a mutational and NMR investigation of pseudoknots derived from the bacteriophage T2 gene 32 mRNA and the retroviral Gag–Pro frameshift site. *RNA* **8**, 981–996 (2002).
78. Cornish, P. V., Hennig, M. & Giedroc, D. P. A loop 2 cytidine-stem 1 minor groove interaction as a positive determinant for pseudoknot-stimulated-1 ribosomal frameshifting. *Proc. Natl Acad. Sci. USA* **102**, 12694–12699 (2005).
79. Hansen, T. M., Reihani, S. N., Oddershede, L. B. & Sorensen, M. A. Correlation between mechanical strength of messenger RNA pseudoknots and ribosomal frameshifting. *Proc. Natl Acad. Sci. USA* **104**, 5830–5835 (2007).
80. Shigemoto, K. *et al.* Identification and characterisation of a developmentally regulated mammalian gene that utilises –1 programmed ribosomal frameshifting. *Nucleic Acids Res.* **29**, 4079–4088 (2001).
81. Manktelow, E., Shigemoto, K. & Brierley, I. Characterization of the frameshift signal of Edr, a mammalian example of programmed –1 ribosomal frameshifting. *Nucleic Acids Res.* **33**, 1553–1563 (2005).
82. Wills, N. M., Moore, B., Hammer, A., Gesteland, R. F. & Atkins, J. F. A functional-1 ribosomal frameshift signal in the human paraneoplastic *Ma3* gene. *J. Biol. Chem.* **281**, 7082–7088 (2006).
83. Jacobs, J. L., Belew, A. T., Rakauskaite, R. & Dinman, J. D. Identification of functional, endogenous programmed-1 ribosomal frameshift signals in the genome of *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **35**, 165–174 (2007).
84. Nonin-Lecomte, S., Felden, B. & Dardel, F. NMR structure of the *Aquifex aeolicus* tmRNA pseudoknot PK1: new insights into the recoding event of the ribosomal trans-translation. *Nucleic Acids Res.* **34**, 1847–1853 (2006).
85. Kaur, S., Gillet, R., Li, W., Gursky, R. & Frank, J. Cryo-EM visualization of transfer messenger RNA with two SmpBs in a stalled ribosome. *Proc. Natl Acad. Sci. USA* **103**, 16484–16489 (2006).
86. Moore, S. D. & Sauer, R. T. The tmRNA system for translational surveillance and ribosome rescue. *Annu. Rev. Biochem.* **76**, 101–124 (2007).
87. Ivanov, I. P. & Atkins, J. F. Ribosomal frameshifting in decoding antizyme mRNAs from yeast and protists to humans: close to 300 cases reveal remarkable diversity despite underlying conservation. *Nucleic Acids Res.* **35**, 1842–1858 (2007).
88. Zimmer, M., Sattelberger, E., Inman, R. B., Calendar, R. & Loessner, M. J. Genome and proteome of *Listeria monocytogenes* phage PSA: an unusual case for programmed +1 translational frameshifting in structural protein synthesis. *Mol. Microbiol.* **50**, 303–317 (2003).
89. Wills, N. M., Gesteland, R. F. & Atkins, J. F. Evidence that a downstream pseudoknot is required for translational read-through of the Moloney murine leukemia virus gag stop codon. *Proc. Natl Acad. Sci. USA* **88**, 6991–6995 (1991).
90. Feng, Y. X., Yuan, H., Rein, A. & Levin, J. G. Bipartite signal for read-through suppression in murine leukemia virus mRNA: an eight-nucleotide purine-rich sequence immediately downstream of the gag termination codon followed by an RNA pseudoknot. *J. Virol.* **66**, 5127–5132 (1992).
91. Wills, N. M., Gesteland, R. F. & Atkins, J. F. Pseudoknot-dependent read-through of retroviral gag termination codons: importance of sequences in the spacer and loop 2. *EMBO J.* **13**, 4137–4144 (1994).
92. Alam, S. L., Wills, N. M., Ingram, J. A., Atkins, J. F. & Gesteland, R. F. Structural studies of the RNA pseudoknot required for readthrough of the gag-termination codon of murine leukemia virus. *J. Mol. Biol.* **288**, 837–852 (1999).
93. Gluck, T. C., Wills, N. M., Gesteland, R. F. & Draper, D. E. Folding of an mRNA pseudoknot required for stop codon readthrough: effects of mono- and divalent ions on stability. *Biochemistry* **36**, 16173–16186 (1997).
94. Orlova, M., Yueh, A., Leung, J. & Goff, S. P. Reverse transcriptase of Moloney murine leukemia virus binds to eukaryotic release factor 1 to modulate suppression of translational termination. *Cell* **115**, 319–331 (2003).
95. Pinck, M., Yot, P., Chapeville, F. & Duranton, H. M. Enzymatic binding of valine to the 3' end of TYMV-RNA. *Nature* **226**, 954–956 (1970).
96. Giege, R. Interplay of tRNA-like structures from plant viral RNAs with partners of the translation and replication machineries. *Proc. Natl Acad. Sci. USA* **93**, 12078–12081 (1996).
97. Matsuda, D. & Dreher, T. W. The tRNA-like structure of turnip yellow mosaic virus RNA is a 3'-translational enhancer. *Virology* **321**, 36–46 (2004).
- A detailed investigation of the role of the TYMV 3' UTR in protein synthesis.**
98. Barends, S., Bink, H. H., van den Worm, S. H., Pleij, C. W. & Kraal, B. Entrapping ribosomes for viral translation: tRNA mimicry as a molecular Trojan horse. *Cell* **112**, 123–129 (2003).

99. Matsuda, D. & Dreher, T. W. Cap- and initiator tRNA-dependent initiation of TYMV polyprotein synthesis by ribosomes: evaluation of the Trojan horse model for TYMV RNA translation. *RNA* **13**, 129–137 (2007).
100. Kolakofsky, D. & Weissmann, C. Possible mechanism for transition of viral RNA from polysome to replication complex. *Nature New Biol.* **231**, 42–46 (1971).
101. Gamarnik, A. V. & Andino, R. Switch from translation to RNA replication in a positive-stranded RNA virus. *Genes Dev.* **12**, 2295–2304 (1998).
102. Barton, D. J., Morasco, B. J. & Flanagan, J. B. Translating ribosomes inhibit poliovirus negative-strand RNA synthesis. *J. Virol.* **73**, 10104–10112 (1999).
103. Deiman, B. A., Koenen, A. K., Verlaan, P. W. & Pleij, C. W. Minimal template requirements for initiation of minus-strand synthesis *in vitro* by the RNA-dependent RNA polymerase of turnip yellow mosaic virus. *J. Virol.* **72**, 3965–3972 (1998).
104. Singh, R. N. & Dreher, T. W. Specific site selection in RNA resulting from a combination of nonspecific secondary structure and –CCR– boxes: initiation of minus strand synthesis by turnip yellow mosaic virus RNA-dependent RNA polymerase. *RNA* **4**, 1083–1095 (1998).
105. Matsuda, D., Yoshinari, S. & Dreher, T. W. eEF1A binding to aminoacylated viral RNA represses minus strand synthesis by TYMV RNA-dependent RNA polymerase. *Virology* **321**, 47–56 (2004).
106. Choi, Y. G. & Rao, A. L. Packaging of brome mosaic virus RNA3 is mediated through a bipartite signal. *J. Virol.* **77**, 9750–9757 (2003).
107. Koenig, R. *et al.* Nemesis ring necrosis virus: a new tymovirus with a genomic RNA having a histidylatable tobamovirus-like 3' end. *J. Gen. Virol.* **86**, 1827–1833 (2005).
108. van Belkum, A., Abrahams, J. P., Pleij, C. W. & Bosch, L. Five pseudoknots are present at the 204 nucleotides long 5' noncoding region of tobacco mosaic virus RNA. *Nucleic Acids Res.* **13**, 7673–7686 (1985).
109. Felden, B., Florentz, C., Giege, R. & Westhof, E. A central pseudoknotted three-way junction imposes tRNA-like mimicry and the orientation of three 5' upstream pseudoknots in the 3' terminus of tobacco mosaic virus RNA. *RNA* **2**, 201–212 (1996).
110. Leathers, V., Tanguay, R., Kobayashi, M. & Gallie, D. R. A phylogenetically conserved sequence within viral 3' untranslated RNA pseudoknots regulates translation. *Mol. Cell Biol.* **13**, 5331–5347 (1993).
111. Zeenko, V. V. *et al.* Eukaryotic elongation factor 1A interacts with the upstream pseudoknot domain in the 3' untranslated region of tobacco mosaic virus RNA. *J. Virol.* **76**, 5678–5691 (2002).
112. Takamatsu, N., Watanabe, Y., Meshi, T. & Okada, Y. Mutational analysis of the pseudoknot region in the 3' noncoding region of tobacco mosaic virus RNA. *J. Virol.* **64**, 3686–3693 (1990).
113. Carr, T. *et al.* Tobamovirus infection is independent of HSP101 mRNA induction and protein expression. *Virus Res.* **121**, 33–41 (2006).
114. Ray, D., Na, H. & White, K. A. Structural properties of a multifunctional T-shaped RNA domain that mediate efficient tomato bushy stunt virus RNA replication. *J. Virol.* **78**, 10490–10500 (2004).
115. Fabian, M. R. & White, K. A. 5'-3' RNA–RNA interaction facilitates cap- and poly(A) tail-independent translation of tomato bushy stunt virus mRNA: a potential common mechanism for tombusviridae. *J. Biol. Chem.* **279**, 28862–28872 (2004).
116. Fabian, M. R. & White, K. A. Analysis of a 3'-translation enhancer in a tombusvirus: a dynamic model for RNA–RNA interactions of mRNA termini. *RNA* **12**, 1304–1314 (2006).
117. Na, H., Fabian, M. R. & White, K. A. Conformational organization of the 3' untranslated region in the tomato bushy stunt virus genome. *RNA* **12**, 2199–2210 (2006).
118. Ferre-D'Amare, A. R., Zhou, K. & Doudna, J. A. Crystal structure of a hepatitis delta virus ribozyme. *Nature* **395**, 567–574 (1998).
- The crystal structure of a ribozyme revealed how the pseudoknot has a key role in folding and function.**
119. Salehi-Ashtiani, K., Luptak, A., Litovchick, A. & Szostak, J. W. A genome-wide search for ribozymes reveals an HDV-like sequence in the human CPEB3 gene. *Science* **313**, 1788–1792 (2006).
120. Klein, D. J. & Ferre-D'Amare, A. R. Structural basis of *glmS* ribozyme activation by glucosamine-6-phosphate. *Science* **313**, 1752–1756 (2006).
121. Cochran, J. C., Lipchok, S. V. & Strobel, S. A. Structural investigation of the *GlmS* ribozyme bound to its catalytic cofactor. *Chem. Biol.* **14**, 97–105 (2007).
122. Chen, J. L. & Greider, C. W. An emerging consensus for telomerase RNA structure. *Proc. Natl Acad. Sci. USA* **101**, 14683–14684 (2004).
123. Autexier, C., Pruzan, R., Funk, W. D. & Greider, C. W. Reconstitution of human telomerase activity and identification of a minimal functional region of the human telomerase RNA. *EMBO J.* **15**, 5928–5935 (1996).
124. Comolli, L. R., Smirnov, I., Xu, L., Blackburn, E. H. & James, T. L. A molecular switch underlies a human telomerase disease. *Proc. Natl Acad. Sci. USA* **99**, 16998–17003 (2002).
125. Theimer, C. A., Finger, L. D., Trantirek, L. & Feigon, J. Mutations linked to dyskeratosis congenita cause changes in the structural equilibrium in telomerase RNA. *Proc. Natl Acad. Sci. USA* **100**, 449–454 (2003).
126. Yingling, Y. G. & Shapiro, B. A. The prediction of the wild-type telomerase RNA pseudoknot structure and the pivotal role of the bulge in its formation. *J. Mol. Graph Model* **25**, 261–274 (2006).
127. Kim, N. W. *et al.* Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011–2015 (1994).
128. Chen, Z. *et al.* Telomerase activity in Kaposi's sarcoma, squamous cell carcinoma, and basal cell carcinoma. *Exp. Biol. Med.* **226**, 753–757 (2001).
129. Fragnet, L., Blasco, M. A., Klapper, W. & Rasschaert, D. The RNA subunit of telomerase is encoded by Marek's disease virus. *J. Virol.* **77**, 5985–5996 (2003).
130. Trapp, S. *et al.* A virus-encoded telomerase RNA promotes malignant T cell lymphomagenesis. *J. Exp. Med.* **203**, 1307–1317 (2006).
131. Artandi, S. E. Telomerase flies the coop: the telomerase RNA component as a viral-encoded oncogene. *J. Exp. Med.* **203**, 1143–1145 (2006).
132. Wyatt, J. R., Puglisi, J. D. & Tinoco, I. Jr. RNA pseudoknots. Stability and loop size requirements. *J. Mol. Biol.* **214**, 455–470 (1990).
133. Klovin, J. & van Duin, J. A long-range pseudoknot in Q $\beta$  RNA is essential for replication. *J. Mol. Biol.* **294**, 875–884 (1999).
134. Paul, C. P., Barry, J. K., Dinesh-Kumar, S. P., Brault, V. & Miller, W. A. A sequence required for -1 ribosomal frameshifting located four kilobases downstream of the frameshift site. *J. Mol. Biol.* **310**, 987–999 (2001).
135. Barry, J. K. & Miller, W. A. A-1 ribosomal frameshift element that requires base pairing across four kilobases suggests a mechanism of regulating ribosome and replicase traffic on a viral RNA. *Proc. Natl Acad. Sci. USA* **99**, 11133–11138 (2002).
- This paper provides an elegant description of how long-range interactions in viral RNAs can be exploited for regulatory purposes.**
136. Neuman, B. W. *et al.* Inhibition, escape and attenuated growth of severe acute respiratory syndrome coronavirus treated with antisense morpholino oligomers. *J. Virol.* **79**, 9665–9676 (2005).
137. Mathews, D. H. & Turner, D. H. Prediction of RNA secondary structure by free energy minimization. *Curr. Opin. Struct. Biol.* **16**, 270–278 (2006).
138. Reeder, J. & Giegerich, R. Design, implementation and evaluation of a practical pseudoknot folding algorithm based on thermodynamics. *BMC Bioinformatics* **5**, 104 (2004).
139. Lyngso, R. B. & Pedersen, C. N. RNA pseudoknot prediction in energy-based models. *J. Comput. Biol.* **7**, 409–427 (2000).
140. Shapiro, B. A., Yingling, Y. G., Kasprzak, W. & Bindewald, E. Bridging the gap in RNA structure prediction. *Curr. Opin. Struct. Biol.* **17**, 157–165 (2007).
141. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505–510 (1990).
142. Tuerk, C., MacDougall, S. & Gold, L. RNA pseudoknots that inhibit human immunodeficiency virus type 1 reverse transcriptase. *Proc. Natl Acad. Sci. USA* **89**, 6988–6992 (1992).
143. Kensch, O. *et al.* HIV-1 reverse transcriptase-pseudoknot RNA aptamer interaction has a binding affinity in the low picomolar range coupled with high specificity. *J. Biol. Chem.* **275**, 18271–18278 (2000).
144. Chaloin, L., Lehmann, M. J., Szczakiel, G. & Restle, T. Endogenous expression of a high-affinity pseudoknot RNA aptamer suppresses replication of HIV-1. *Nucleic Acids Res.* **30**, 4001–4008 (2002).
145. Held, D. M., Kissel, J. D., Saran, D., Michalowski, D. & Burke, D. H. Differential susceptibility of HIV-1 reverse transcriptase to inhibition by RNA aptamers in enzymatic reactions monitoring specific steps during genome replication. *J. Biol. Chem.* **281**, 25712–25722 (2006).
146. Jaeger, J., Restle, T. & Steitz, T. A. The structure of HIV-1 reverse transcriptase complexed with an RNA pseudoknot inhibitor. *EMBO J.* **17**, 4535–4542 (1998).
147. Sayer, N., Ibrahim, J., Turner, K., Tahiri-Alaoui, A. & James, W. Structural characterization of a 2'-F-RNA aptamer that binds a HIV-1 SU glycoprotein, gp120. *Biochem. Biophys. Res. Commun.* **293**, 924–931 (2002).
148. Kim, M. Y. & Jeong, S. RNA aptamers that bind the nucleocapsid protein contain pseudoknots. *Mol. Cell* **16**, 413–417 (2003).
149. Noller, H. F. RNA structure: reading the ribosome. *Science* **309**, 1508–1514 (2005).

### Acknowledgements

We are grateful to Stephen Smerdon for critical reading of the manuscript.

### Competing interests statement

The authors declare no competing financial interests.

### DATABASES

The following terms in this article are linked online to:

Entrez Genome Project:  
<http://www.ncbi.nlm.nih.gov/sites/entrez>  
*Escherichia coli* | *Listeria monocytogenes*  
 Entrez Genome: <http://www.ncbi.nlm.nih.gov/sites/entrez?cmd=search&db=pubmed>  
 BWYV | CrPV | HCV | HDV | PSIV | SARS-CoV | TBSV | TMV | TYMV

### FURTHER INFORMATION

Ian Brierley's homepage:  
<http://www.path.cam.ac.uk/pages/brierley>  
 Robert Gilbert's homepage:  
<http://www.strubi.ox.ac.uk/people/gilbert>  
 RNAmotif:  
<http://www.scripps.edu/mb/case/casegr-sh-3.5.html>  
 PknotsRG: <http://bibiserv.techfak.uni-bielefeld.de/download/tools/pknotsrg.html>  
 HotKnots:  
<http://www.cs.ubc.ca/labs/beta/Software/HotKnots>  
 HPKNOTTER:  
<http://bioalgorithm.life.nctu.edu.tw/HPKNOTTER>  
 PseudoViewer: <http://pseudoviewer.inha.ac.kr>  
 PseudoBase:  
<http://biology.leidenuniv.nl/~batenburg/PKBGetS1.html>

### SUPPLEMENTARY INFORMATION

See online article: S1 (table) | S2 (table) | S3 (figure) | S4 (figure) | S5 (figure)

Access to this links box is available online.